

Parameter Estimation of Space-Time Model Using Genetic Algorithm

S. Halim, I. N. Bisono, D. Sunyoto, I. Gendo

Industrial Engineering Department, Universitas Kristen Petra, Surabaya, Indonesia
(halim@petra.ac.id, mlindri@petra.ac.id)

Abstract - The Space-Time Autoregressive Moving-Average (STARMA) model family is a statistical inductive model that can be used to describe stationary (or weak stationary) space-time processes. However, parameter estimation of the model often is not easy to obtain analytically because of the hard computation or the unknown probability density function underlying the data. To ease the difficulty, an approach to estimate the parameter is proposed in this study, i.e. genetic algorithm (GA). GA is one of the meta-heuristic methods widely used in many applications including the parameter estimation. The GA is performed through simulations of various combinations of selection and crossover parameter chromosomes. The estimation, then, was carried out by the help of freeware R. The performance of the GA in estimating parameter is measured in the sense of the minimum residual sum of squares and the Akaike Information Criterion (AIC). In order to have a comparable solution, we employed the STARMA model of assault arrests in 14 districts of Northeast Boston (1969-1974) of Pfeifer and Deutsch. The results show that the performance of the GA is relatively competitive to the classical method. Since GA is simple to apply, it might be considered as one of the alternative methods for estimating space-time model parameters.

Keywords - Space-Time Modeling, STARMA, Genetic Algorithm, Parameter Estimation

I. INTRODUCTION

Many times value of an object depends on time and its surrounding. Home price, pollution level, and criminal levels are some examples of space-time variables. Modeling such variables using time series only might be inadequate as the variables might contain spatial pattern. The Space-Time Auto Regressive Moving-Average (STARMA) models have gained widespread popularity in many domains, including imaging, transport, business and economics, and hydrology, etc. Anselin, *et al.* [2], LeSage [7]. Inline with those needs, LeSage and Polasek [8], and LeSage and Pace [9] developed *spatial econometrics*. Pfeifer and Deutsch [10] extended Box and Jenkins [2] time series models into *spatial model*.

Three stages procedure for Space-Time modeling presented by Pfeifer and Deutsch [10] consist of model identification, parameter estimation, and model diagnostic check. The parameter estimation was accomplished through maximum likelihood estimates (m.l.e.). In using m.l.e., one need to identify the probability density functions underlying the data, which sometimes is not easy to find. In addition, when we do have the p.d.f., solving the optimal value analytically sometimes is too

intricate. In this study, we propose a heuristic method i.e. genetic algorithm (GA) to estimate STARMA parameters. GA has been proved to be a robust approach on estimating the parameters on the time series and non linear functions simultaneously [11]. As a study case and in order to have a comparable solution, we employed the STARMA model of assault arrests in 14 districts of Northeast Boston (1969-1974) of Pfeifer and Deutsch [10].

The paper is organized as follows. Section 2 introduces the methods. The GA and parameter estimation is discussed in Section 3. At last, Section 4 provides conclusion.

II. METHODOLOGY

A. The Space-Time Autoregressive Moving Average Model

When N observable time series, $z(t)$, contain pattern of regions, the model should represent the dependence among observations at neighboring regions. Models that explicitly attempt to explain these dependencies across space are referred to as space-time models. Space-time models are characterized by autoregressive and moving average terms lagged in both time and space (Pfeifer and Deutsch, [10]). N regions in a space could represent variation situations such as districts of a town or cities of a province.

The general model of STARMA is given as follows,

$$z(t) = \sum_{k=1}^p \sum_{l=0}^{\lambda_k} \phi_{kl} W^{(l)} z(t-k) - \sum_{k=1}^q \sum_{l=0}^{m_k} \theta_{kl} W^{(l)} \varepsilon(t-k) + \varepsilon(t) \quad (1)$$

where p is the autoregressive order, q is the moving average order, λ_k is the spatial order of the k -th autoregressive term, m_k is the spatial order of the k -th moving average term, ϕ_{kl} and θ_{kl} are parameters, $W^{(l)}$ is the weight matrix, and $\varepsilon(t)$ is the random normal error. The weight $W^{(l)}$ is defined based on physical properties of the system, for instance border length of two neighbor cities, distance from city center, natural border such as river or mountain, and level of accessibility from city i to city j . Weights are given in accordance with hierarchical ordering of spatial neighbor. The first order neighbors are pointed to the nearest sites. The second order neighbors are for sites located further than the first order neighbors but closer to the third order neighbors.

The identification process, the first stage of building STARMA model, includes selection one of alternative models (AR, MA, or ARMA) that represent the observed data. The model selection is rooted in the behavior of autocorrelation functions (ACF) and partial auto-

correlation functions (PACF). The model might be STAR whenever the ACF tails off. While whenever both ACF and PACF tail off, it might be STARMA.

In the second stage, parameter estimation, Pfeifer and Deutsch [10] obtained the best parameter estimates for Φ and θ using maximum likelihood estimates,

$$f(\varepsilon | \Phi, \theta \sigma^2) = (2\pi)^{-TN/2} (\sigma^2)^{-TN/2} \exp\left(-\frac{S(\Phi, \theta)}{2\sigma^2}\right) \quad (2)$$

$$\text{where } S(\Phi, \theta) = \varepsilon' \varepsilon = \sum_{i=1}^N \sum_{t=1}^T \varepsilon_i(t)^2. \quad (3)$$

$\varepsilon(t)$ can be recursively calculated from $z(t)$ using the following equation:

$$\varepsilon(t) = z(t) - \sum_{k=1}^p \sum_{l=0}^{\lambda_k} \phi_{kl} W^{(l)} z(t-k) + \sum_{k=1}^q \sum_{l=0}^{m_k} \theta_{kl} W^{(l)} \varepsilon(t-k) \quad (4)$$

for $t = 1, 2, \dots, T$.

In the calculation process, Pfeifer and Deutsch [10] followed estimation algorithm of Marquardt (1963). The final step, the diagnostic checking, is required to determine whether the model explain the data well. The diagnostic checking comprises two important steps, i.e. residual checking and significant test of estimated parameters. If the model explains the data well, the residuals might be white noise. When the residuals are white noise, then the sample of space-time autocorrelation function should be zero. The appropriate test statistic for hypothesis $\delta_k = 0$ is

$$\frac{(TN - K)[S_*(\hat{\delta}^*) - S_*(\hat{\delta})]}{S_*(\hat{\delta})} \quad (5)$$

which is approximately distributed $F_{1, TN-K}$ under the null hypothesis. Any estimated parameter which is not statistically significant will be withdrawn from the model, thus a simpler model will be considered as a new candidate.

B. Genetic Algorithm

Genetic algorithm (GA) is a meta-heuristic method used to find solution based on biological evolution process. The process includes crossover natural selection and mutation to obtain an individual with the best gen combination. GA begins with determination of chromosome set (solution set) in terms of binary (1 and 0). Next, the selection is conducted based on the fitness value. The chromosome with the highest fitness value is retained while the rest are removed. The selected chromosomes, then, experience reproduction process to be parents. Haupt and Haupt [6] categorized five types of selection methods namely from top to bottom, random pairing, tournament, roulette wheel rank weighting, and roulette wheel cost weighting.

The next generation, the off springs, are created using crossover and mutation [4]. Haupt and Haupt [6] pointed three types of crossover that are generally used, i.e. single-point crossover, double-point crossover, and

uniform crossover. The mutation employed is a single point mutation which changes 1 (one) in binary into 0 (zero) and vice versa, 0 into 1. However, mutation does not count for elite solution because chromosome with the best fitness does not need to be changed. It is called elitism [6]. GA parameters that would be applied to estimate parameters of STARMA model followed Grenfenstette setting [5]; the population size equals 80, the crossover rate equals 0.45, and the mutation rate equals 0.01. The generation number was determined to be 100 generation.

C. Akaike Information Criterion (AIC)

AIC could be utilized to compare competing models. A model with minimum AIC will be considered as the best model as it has minimum distance to the truth. Based on normality and independent error assumptions, Akaike [1] formulated the AIC as follows,

$$AIC = 2k + n[\ln(2\pi RSS / n) + 1] \quad (6)$$

where k is number of parameters, n is number of observations and RSS is the residual sum of squares.

III. RESULTS

A. Data and Variable

Data of criminal level in *Northeast Boston* will be used as a case study. The increasing or decreasing level of crime in one region might be affected by the level of crime in its neighbor regions. Usually, the criminal will committed crime in the neighborhood area. The socio-economic condition might influence the level of crime while the socio-economic condition in neighborhood regions are dependent. The criminal level in one region might also be influenced by the previous level of crime. Therefore, the criminal data consist of spatial and temporal pattern and should be modeled using space-time modeling. Fig. 1 shows the 14 Northeast Boston districts where the investigated criminal level occurred on interval 6 years (1969-1974). Plot of the number of criminal cases per month can be seen in Fig. 2. The definition of spatial order neighbors for the 14 districts follows Pfeifer and Deutsch [10].

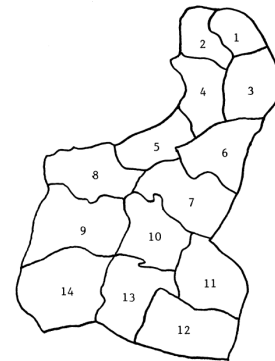


Fig. 1. The map of Northeast Boston districts

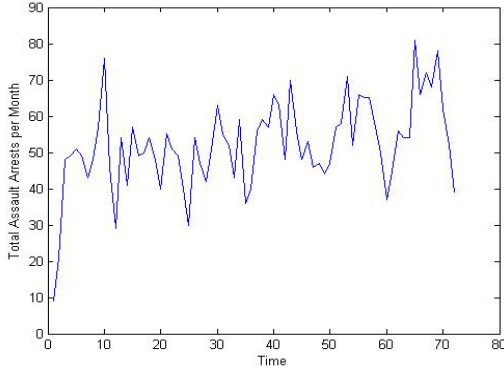


Fig. 2. Number of criminal cases per month in Northeast Boston

B. Parameter Estimation of STARMA using Genetic Algorithm

Often it is not easy to estimate parameters analytically because the distribution might be unknown; even when the distribution is known sometimes the differentiation or the solution of the differential equation is difficult to obtain. Therefore, we proposed a genetic algorithm as an alternative solution in this case. Solution of parameter estimation is obtained by means of Statistical freeware R.2.9.0. Fig. 3 shows the scheme of genetic algorithm used on parameter estimation of the STARMA model.

The first step on the parameter estimation of STARMA model is to identify the spatial lags, P , Q , and time lags, OP , OQ . The time lags, as usual, are identified using the classical ACF and PACF [3], while the spatial lags are identified using the spatial order system [10]. We also need an upper bound (UP), lower bound (LB), and δ as inputs. UP and LB provide parameter investigation range, while δ determine the degree of accuracy of the investigation. If we want the accuracy up to four decimals then δ should be 10^{-4} . These three values determine number of bits (N_{bits}) in one chromosome as it is stated in the following inequality

$$2^{N_{bits}-1} < (UP - LB) \cdot 10^n \leq 2^{N_{bits}} - 1 \quad (7)$$

where n is the degree of accuracy [4].

In building the STARMA model, the series is assumed as stationary. Therefore, the parameters space should be set on a unit ball. In this study the values we used are 1 and -1 for UP and LB respectively. When we chose four decimal (10^{-4}) for δ , the space of each parameter will have

$$\frac{UP - LB}{\delta} = \frac{1 - (-1)}{10^{-4}} = 20000 \text{ possible values.}$$

The binary number of 20000 is 11111010000. The latter contains 15 digits, so the number of bits per chromosome should be 15 as well. The 15 digits binary value of each chromosome will be generated randomly and then transformed to decimal. The conversion from binary into decimal will result in discrete number, so as quantization level is carried out using the following. To convert the following equations [6];

$$Q = \frac{1}{(2^{-1} + 2^{-2} + \dots + 2^{-N_{bits}})} \times quant \quad (8)$$

where, $quant$ = quantization vector = $[2^{-1} \ 2^{-2} \ \dots \ 2^{-N_{bits}}]$.

$G = [b_1, b_2, \dots, b_{N_{bits}}]$, b is binary number.

$$P_{quant} = GQ^T \quad (9)$$

$$q_n = P_{quant}(UP - LB) + LB \quad (10)$$

Based on (8) the quantization vector is $Q = [0.5 \ 0.25 \ 0.125 \ 0.0625 \ 0.0313 \ 0.0156 \ 0.0078 \ 0.0039 \ 0.002 \ 0.001 \ 0.0005 \ 0.0002 \ 0.0001 \ 0.0001 \ 0.0000]$.

Now using (9) we transform the chromosome to a real number. We then adjust the P_{quant} by applying (10), such that its value will fall between the upper and lower bound, i.e., -1 and 1. Take for example; a chromosome selected randomly is 110110011101101, then P_{quant} equal to 0.8510. The parameter value become

$$q_n = 0.8510(1 - (-1)) + (-1) = 0.7020.$$

Each parameter value is obtained following the same procedure.

The general steps taken in parameter estimation of STARMA using GA are explained as follows;

- Initial population is determined by randomly generate k chromosomes for each parameter i.e. Φ (ϕ) and θ (θ).
- The number of chromosomes that will be kept for the next generation, N_{keep} , are based on fitness value. Residual sum of squares (RSS) from fitness function (3) and (4) is used as fitness value. Next, the initial population is ascendingly ordered, from the smallest to the highest RSS. Φ dan θ chromosomes in $N_{keep}+1$ up to k position are thrown away and replaced with new individuals.
- Selection process. This process is used to determine two chromosomes to be the parents. There are five selection processes employed in this paper, namely *pairing from top to bottom*, *random pairing*, *tournament*, *roulette wheel rank weighting*, and *roulette wheel cost weighting*.
- Crossover. Two new individuals were born as a result of crossover of two parents. The three crossover methods used are *single-point crossover*, *double-point crossover*, and *uniform crossover*.
- Mutation. Φ dan θ chromosomes with the smallest RSS (first position) will not get mutation because of elitism.
- Population evaluation after regeneration process will be ended up similarly as in step (b).

Repeated step (c) to (f) to get the next generation as many as number of generations needed. Parameter estimation process has finished whenever the regeneration process finished. The parameter values to be used in the model are ϕ and θ values with the smallest RSS (first position) in the last population. The smallest RSS of all combination will then be used to calculate the AIC (6).

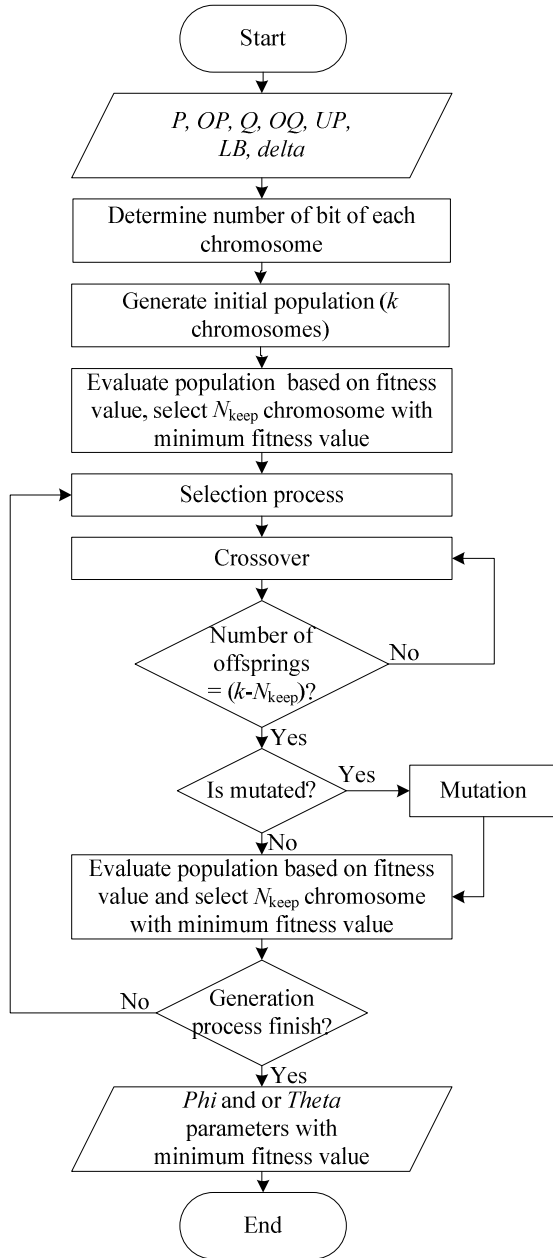


Fig. 3. Flowchart of STARMA model parameter estimation using Genetic Algorithm

TABLE I
COMPARISON OF PARAMETER ESTIMATION VALUES BETWEEN CLASSICAL AND GENETIC ALGORITHM

	STARIMA (0, 1, 1 ₀)		STARIMA (0, 1, 1 ₁)	
	Classical Approach	Genetic Algorithm	Classical Approach	Genetic Algorithm
$\hat{\theta}_{10}$	0.8030	0.7924	0.8120	0.8080
$\hat{\theta}_{11}$	-	-	-0.0920	-0.0972
RSS	6504.6790	6534.0294	6457.2660	6478.4143
AIC	4692.5184	4694.9932	4685.2465	4688.4965

IV. DISCUSSION

Parameter estimation was conducted by using all combination selection methods and crossover. We run the program ten times to achieve the minimum AIC and RSS for each combination. The simulation results are available in Table 1. It can be seen that the estimation using genetic algorithm might not perform better than classical algorithm. However, the AIC values of the two do not differ significantly.

In this study, we employed 15 combination operator methods in genetic algorithm for parameter estimation of STARMA model. The performance of each combination is examined through the rate of each combination to attain minimum RSS. The parameters in *genetic algorithm* used were following Grenfenstette *setting*, namely *crossover rate* 0.45, *mutation rate* 0.01, and 100 generations. Population size was varied from 10 to 80 chromosomes. Plots of fitness value as a function of generation number for population size 10 and 80 are shown in Fig. 4 and 5.

The analysis of all experiments show that population size does not affect the minimum RSS. It can be seen that all combination with various population size can achieve minimum RSS in relatively small variation so as considered as the same.

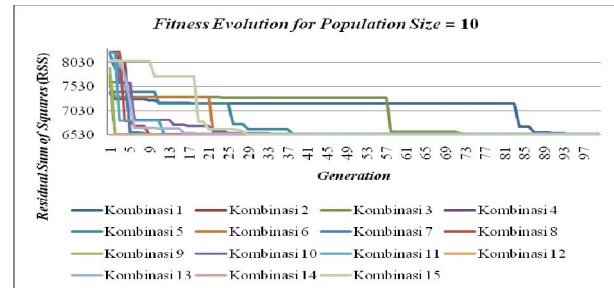


Fig. 4. Behavior of RSS of each combination with *Population Size* 10

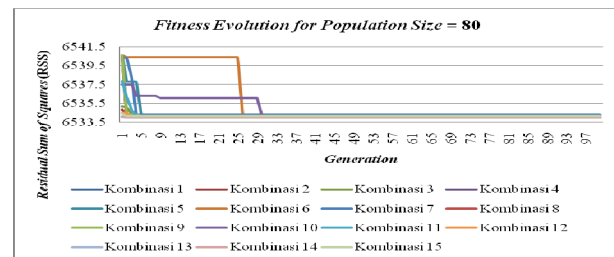


Fig. 5. Behavior of RSS of each combination with *Population Size* 80

V. CONCLUSION

In GA application, the number of generation and population size should be not too small in order to reduce the random effect and stabilize the result. The simulation results show that number of generation 100 and population size 80 have been adequate to attain good results. As a conclusion, GA performed competitively to classical method to estimate model parameters. Thus, GA can be considered as an alternative method in *space-time modeling* parameter estimation.

ACKNOWLEDGMENT

The authors are very grateful to Prof. Pfeifer for providing us the data and to the reviewers for their valuable comments and suggestions.

REFERENCES

- [1] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.
- [2] L. Anselin, R. J. Florax, and S. J. Rey, *Advances in Spatial Econometrics: Methodology, Tools and Applications*. Heidelberg: Springer, 2004.
- [3] G. E. P. Box, and G. M. Jenkins, *Time series analysis, forecasting and control*. San Francisco: Holden-Day, 1994.
- [4] M. Gen, and R. Cheng, *Genetic algorithms and engineering design*. New York: John Wiley & Sons, Inc., 1997.
- [5] J. Greffenstette, "Optimization of control parameters for genetic algorithms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 16, no. 1, pp. 122–128, 1986.
- [6] R. J. Haupt, and S. E. Haupt, *Practical genetic algorithms* (2nd ed.). New York: John Wiley & Sons, Inc., 2004.
- [7] J. P. LeSage, "Econometric toolbox for matlab." <http://www.spatial-econometrics.com/>, Retrieved September 16th, 2008.
- [8] J. P. LeSage, and W. Polasek, *Incorporating transportation network structure in spatial econometric models of commodity-flows*. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=924613, Retrieved September 20th, 2008.
- [9] J. P. LeSage, and R. K. Pace, *Spatial econometric modeling of origin – destination flows*. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=924609. Retrieved September 20th, 2008.
- [10] P. E. Pfeifer, and S. J. Deutsch, "A three-stage iterative procedure for space-time modeling," *Technometrics*, vol. 22, pp. 35–47, 1980.
- [11] A. F. Sheta, and K. De Jong, "Time-series forecasting using GA-tuned radial basis functions," *Information Sciences*, vol. 33, pp. 221–228, 2001.